FROM SPEED PROFILE DATA TO ANALYSIS OF BEHAVIOUR

– Classification by Pattern Recognition Techniques –

Aliaksei LAURESHYN
Traffic and Roads, Department of Technology and Society
Faculty of Engineering LTH
Lund University,
Lund, Sweden

Kalle ÅSTRÖM
Centre for Mathematical Sciences
Faculty of Engineering LTH
Lund University
Lund, Sweden

Karin BRUNDELL-FREIJ
Analysis & Strategy
WSP Group Sweden
Malmo, Sweden

(Received February 18, 2009)

Classification of speed profiles is necessary to allow interpretation of automatic speed measurements in terms of road user behaviour. Aggregation without considering variation in individual profile shapes easily leads to aggregation bias, while classification based on exogenous criteria runs the risk of loosing important information on behavioural (co-) variation. In this paper we test how three pattern recognition techniques (cluster analysis, supervised learning and dimension reduction) can be applied to automatically classify the shapes of speed profiles of individual vehicles into interpretable types, with a minimum of *a priori* assumptions. The data for the tests is obtained from an automated video analysis system and the results of automated classification are compared to the classification by a human observer done from the video. Normalisation of the speed profiles to a constant number of data points with the same spatial reference allows them to be treated as multidimensional vectors. The *k*-means clustering algorithm groups the vectors (profiles) based on their proximity in multidimensional space. The results are satisfactory, but still the least successful among the tested techniques. Supervised learning (nearest neighbour algorithm tested) uses a training dataset produced beforehand to assign a profile to a specific group. Manual selection of the profiles for the training dataset allows better control of the output results and the classification results are the most successful in the tests. Dimension reduction techniques decrease the amount of data representing each profile by extracting the most typical "features", which allows for better data visualisation and simplifies the classification procedures afterwards. The singular value decomposition (SVD) used in the test performs quite satisfactorily. The general conclusion is that pattern recognition techniques perform well in automated classification of speed profiles compared to classification by a human observer. However, there are no given rules on which technique will perform best.

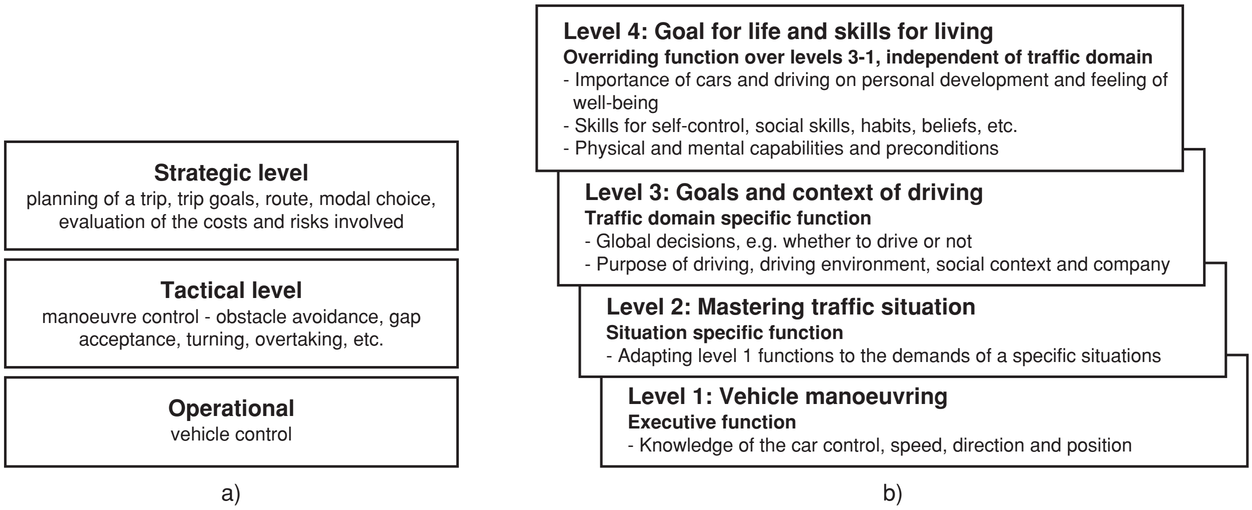
Key Words: Speed profile, Behaviour analysis, Pattern recognition, Clustering, Supervised learning, Dimension reduction

1. INTRODUCTION

The main aim of studying the behaviour in traffic is to understand the reasons for road users to act as they do and learn factors that can affect their actions[1]. One of the main behaviour characteristics is the choice of speed, which has direct implication on safety, efficiency and the size of environmental impacts both for an individual road user and the traffic system in general[2-6]. Not only the general speed level, but also the way the speed changes is of importance. For example, intensive accelerations of a vehicle cause much higher emissions, while braking and the suddenness of braking – jerk – are often regarded as indicators of conflict situations[7-9]. Investigation of the relations between the speed and design of roads and intersections can give an idea of how the optimum speed regime can be reached[10-12].

From the psychological perspective, the speed we observe in any given moment can be seen as a result of the many tasks a road user performs. These can be regarded as being related to decision-making on different levels. For example, Michon[13] describes the following model. The first level (Fig. 1a) is operational and relates to control of the vehicle and decisions about the use of the steering wheel and pedals, gear choice, etc. The second, tactical, level refers to manoeuvring and immediate interactions with other road users. The third, upper level in the hierarchy is strategic and it concerns tasks like trip planning, navigation and route choice. A more recent model, the GADGET-matrix (named after the European project for which it was developed, Fig. 1b), also suggests the fourth level, described as "goals for life and skills for living" and refers to social skills, beliefs, importance of driving for personal well-being and social status, etc.[14].

In view of the described cognitive models, the momentary speed of a road user reflects performance on the



Sources: a) adapted from Human behaviour and traffic safety[13]; b) Gadget-matrix[14]

Fig. 1 The hierarchy of the road user's tasks

lowest operational level. The behaviour, i.e. road user actions in relation to other road users and the road, belongs rather to the second, tactical level and results in a continuous process of speed adaptation to the traffic environment. To understand the behaviour it is necessary to relate the speed changes to the traffic conditions. One of the challenges, however, is to extract the behavioural information from the extensive operational speed data.

One way of collecting the behavioural data is to use qualitative description of the speed changes. An observer makes a note, for example, that a road user "slows down", "stops", "does not alter speed", "yields to another road user", "drives on red", etc. (e.g. Carsten, O.M.J., D.J. Sherborne and J.A. Rothengatter[15]- Martinez, K.L.H. and B.E. Porter[17]). The use of human observers places serious limitations on the amount of data that can be practically collected and also on how detailed the data can be. On the other hand, humans judge the observed situations holistically, and in classifying them might consider dimensions not even captured by the objective measurements of speed or other variables. A typical example of judging traffic situations is traffic conflict studies, where observers are found to be very good in distinguishing between serious and non-serious conflicts and also rating the conflicts' severity. In fact, in an international calibration study there was found a higher agreement between individual observers' judgements than between observers and various objective measures of conflict severity. This only supports the hypothesis that there are relevant subjective dimensions that play a role in the interpretation of observed processes in traffic[18,19].

Another tradition in studying speed data is to use speed loggers installed in vehicles driving in real traffic. After identification of the important characteristics, the data are interpreted in terms of driving patterns used, for example, to produce standard driving cycles for vehicle tests and estimation of the emission factors[9,20,21]. The problem with this method is that the studied population of drivers and vehicles is normally very limited and the information about the traffic situation (e.g. the presence of other road users) is missing.

In recent times, video analysis techniques are becoming a popular tool for traffic data collection[22-24]. This technology provides an opportunity to measure speed with high time frequency and for large populations of road users. However, the problem of interpretation in behaviour-terms still remains, and the greatly increased amount of data that is collected necessitates the analysis to be automated.

Simple aggregation to an average (or 85-percentile) speed profile[12,25] loses information on differences between the individuals and correlation (over time) within individual profiles. It may therefore lead to aggregation bias and misleading final profile shape if the individual profiles are very different in character. Sekine[26] proposes LUNA (Location UNiversal Archive-format) aggregation format, which provides speed distributions at several points along the studied section. This preserves to a greater extent the variation of speeds at each cross-section point, but the longitudinal connections between the points of individual profiles is still lost. None of these approaches utilises the information about the variation in shapes of individual profiles, which can be attributed to different behavioural strategies. If profiles are classified before the aggregation, classification based on exogenous (pre-set) criteria runs a risk of loosing important information on behavioural co-variation.

In order to utilise the advantages of the detailed data contained by large samples of speed profiles, it is necessary to have a method that:

I. differentiates between the behaviour types based on endogenous (derived from data) criteria in a way similar to a human observer;
II. makes use of the systematic variation in the data that can be attributed to different types of behaviour (i.e. analyses shapes of the speed profiles);
III. can handle large amounts of data as produced, for example, by the video analysis techniques.

We suggest using pattern recognition techniques to fill this methodological gap. Pattern recognition is a topic in machine learning theory that aims at classifying data based on *a priori* knowledge or information extracted from the data itself. In this paper we test three techniques (cluster analysis, supervised learning and dimension reduction) according to the criteria I - III described above.

The paper has the following structure. First, we describe the dataset that was used in the tests. Then, we shortly explain the principles of the three pattern recognition techniques one by one and apply them for classification of the speed profiles in the dataset. Finally, the performance of the techniques in the tests is discussed and the conclusions are drawn.

2. THE DATASET: LEFT TURNING VEHICLES AT A SIGNALISED INTERSECTION

The data for the tests was collected as a part of the work on developing a system for automated video analysis[22,27]. A signalised intersection was video-filmed and the system was used to extract the trajectories and speeds of moving road users. Only data on left-turning cars was saved (Fig. 2). The reason for limiting the vehicle type to cars only is that for larger vehicles the accuracy of speed estimated from the video was not very accurate. The left-turning manoeuvre was chosen as it requires interactions with two conflicting flows (on-coming traffic and pedestrians at the pedestrian crossing) and more variation in profile shapes was found compared to other manoeuvres. The speed data were visually checked for consistency and manually corrected in cases of obvious errors in detection. After removing the incomplete profiles (for example if the tracked vehicle was occluded by other vehicles for some time), 253 profiles were left for analysis. The profiles were trimmed and adjusted so that each profile contained 60 data points, evenly distributed along the trajectory between the defined start and end lines. This allowed direct point comparison of the profiles by referring to their order numbers only.

According to the rules, a left-turning vehicle (in a right-hand drive location) must yield both to the traffic coming from the opposite direction and to pedestrians who have green in the same phase. This results in four possible types of situations:

a) There are vehicles coming from the opposite direction, the driver has to yield by braking near the middle line;
b) There is a pedestrian at the pedestrian crossing, the driver has to brake before the crossing;
c) No conflicting traffic is present or the gap is sufficiently large, a turning vehicle proceeds with nearly constant or slightly increasing speed;
d) A driver has to brake both near the middle line and near the pedestrian crossing. This situation is extremely rare, since the pedestrian flow is low and those who are present usually manage to complete

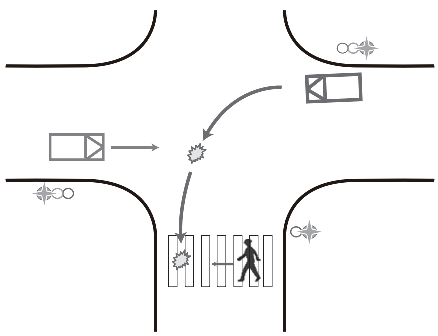


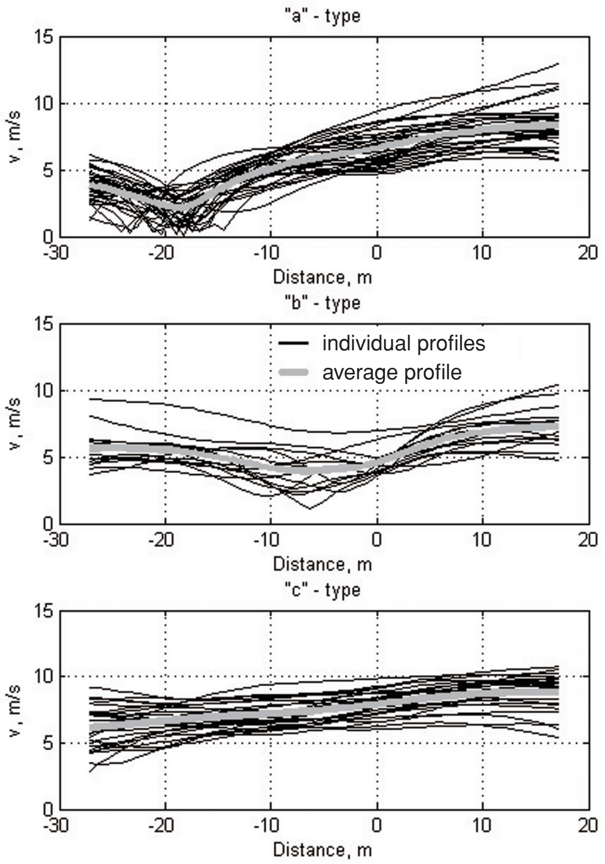
Fig. 2 View of the observation site and conflict points for left turning vehicles

their passage while the driver is waiting at the middle line. There were no such situations in the studied dataset.

The situations in the dataset were classified into three groups (*a*, *b* and *c*- types) by an observer who watched through the video clips from which the speed profiles had been extracted. To avoid influence on the observer's judgements, the observer was not allowed to see the speed profiles corresponding to the situations. The results of the classification are shown on Figure 3.

Examination of the speed profiles revealed that each type of situation has a quite distinctive shape, shown in Figure 4. However, not all the profiles fit the typical shapes perfectly and there is a large group of profiles that appear to be somewhere "in between" two shapes. The observer also expressed difficulties with classifying such situations as the description of more then one type matched them (for example, a car moves slowly forward, avoiding thus sharp baking at the middle line, but still being affected by the on-coming traffic). This is a general problem of the diversity of behaviour forms that complicates its classification. We assume here that the observer's classification is the best possible to achieve and it is used as the "ground truth" in the following tests.

a)



b)

Situation type	Number of situations	Percentage
"a"	*102*	*40%*
"b"	*16*	*6%*
"c"	*135*	*54%*
Total	**253**	**100%**

Fig. 3 Speed profiles in the situations classified by an observer (a) and the distribution of the situations by type (b)

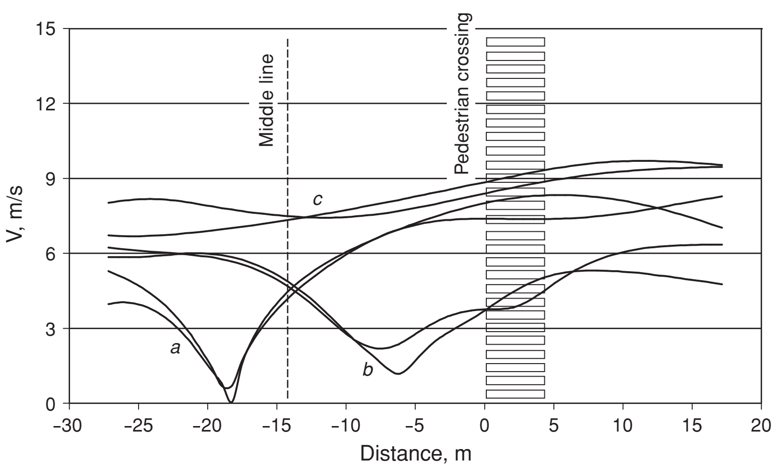


Fig. 4 Three most typical profile shapes: *a* – no on-coming traffic or pedestrians; *b* – driver yields to the on-coming vehicles; *c* – driver yields to pedestrians at the pedestrian crossing

3. PATTERN RECOGNITION TECHNIQUES FOR SPEED PROFILE CLASSIFICATION

3.1 Speed profile as a vector

The concept of the vector, which is defined as a geometrical object having both length and direction, is a key element in linear algebra. A vector in *n*-dimensional space is described by *n* co-ordinates, which is an ordered list of *n* numbers.

All the profiles in the dataset have the same number of points (60, or *n* in the general case) and each point *i* refer to the same location, i.e. the points between the profiles can be directly compared. A possible interpretation of this data is that each profile represents a vector in *n*-dimensional space. The speed at a point *i* is thus the value of the i^{th}-co-ordinate.

This approach helps to better understand the logic of the techniques described in the following sections.

3.2 Cluster analysis

Cluster analysis is a general name for methods of dividing the data into several partitions (clusters) according to some properties considered common for the items within the cluster. Most often this property is proximity, i.e. the items in a cluster are closer to each other or to the cluster centre than to other items or other cluster centres. A clustering algorithm may force the data into a pre-defined number of clusters *k* (*k*-clustering) or find the optimal number of clusters based on the data.

The simplest clustering algorithm is a *k*-means clustering[28]. Figure 5 illustrates its principle in a 2-dimensional example. First, some points are randomly chosen as cluster centres (C_1' and C_2') and each point is assigned to the nearest centre (iteration 1). Then the cluster centres are recalculated as the average of the assigned points and the same procedure is repeated (iteration 2). This is continued until a new reassignment does not differ from the previous one or until the centre's co-ordinates do not change significantly after more iterations. In case of *n* co-ordinates, the algorithm works in the same way, but the distances are calculated according to the rules of *n*-dimensional space.

The advantages of *k*-means are its extreme simplicity and speed of calculation, but the drawbacks are dependence of the results on the choice of initial centres, assumption of the cluster round shapes, etc. More advanced clustering algorithms, such as quality threshold (QT), expectation maximisation (EM) and hierarchical clustering treat many of these problems better, but may incur other problems[29].

Figure 6 illustrates the results of *k*-means clustering of the speed profiles data (to keep diagrams readable only 75 profiles are plotted). Cluster 1 represents mostly the *a*-type profiles. Cluster 2 is a mixture of *b*- and *c*-types and Cluster 3 contains clearly *c*-type profiles. The result is not optimal, and there are several possible ways to improve it. First, a larger number of clusters (for example 9 instead of 3) may be used to make possible differences between the profiles more visible and then merge the clusters belonging to the same type. Another alternative is to repeat cluster analysis within the clusters (in this case it is relevant for Cluster 2), to split the mixed profiles of different types. Comparing clusters 2 and 3 one may note that the profiles in cluster 2 generally have lower speed levels than in Cluster 3. This is not surprising since the *k*-means algorithm clusters the profiles based on the distance between the points, i.e. profiles with different shapes close to each other have a higher chance of appearing in the same cluster than profiles having the same shape but different speed levels. This may be tackled by, for example, using the derivative of speed as a co-ordinate at each point instead of speed itself, since it will produce more

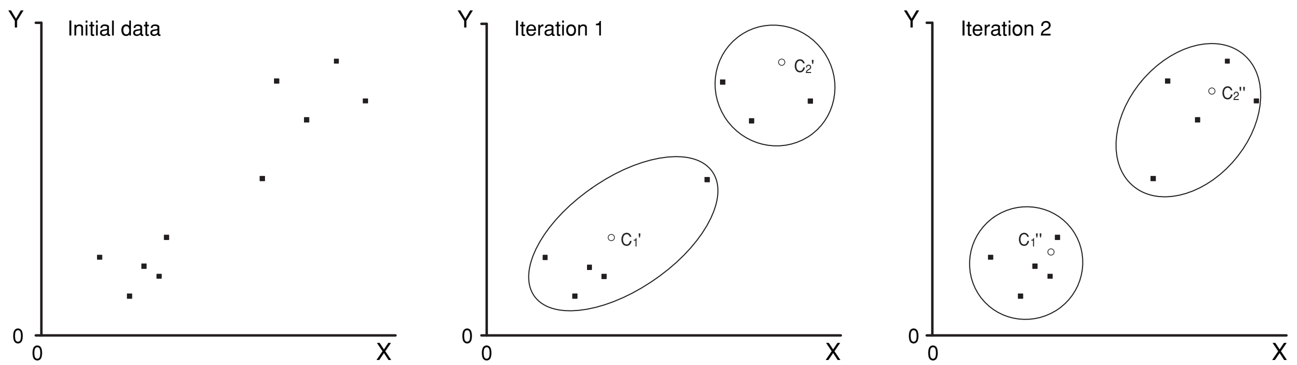


Fig. 5 The *k*-means clustering algorithm at work

level difference between the profiles of different shapes.

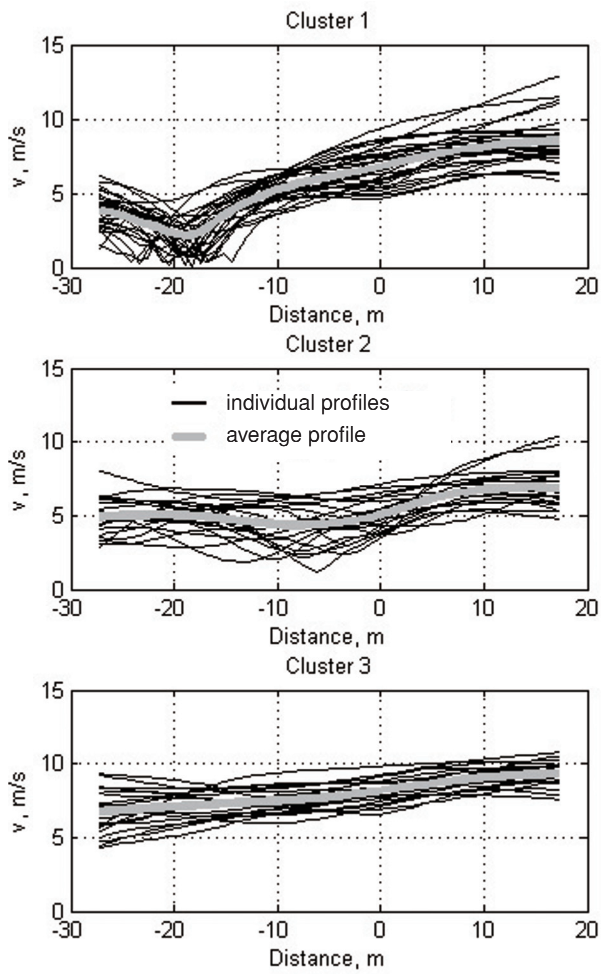
3.3 Supervised Learning

Supervised learning is another technique that may be used for data classification. The main difference from clustering is, however, that the classification function is learnt from a training dataset containing both the input objects and the desired outputs. There is a wide range of classification algorithms developed, where *k*-nearest neighbours is one of the simplest. Figure 7 illustrates how the algorithm works in a two-dimensional example. The training dataset contains objects of two types – black squares and white circles. In order to decide to which type a new object belongs, the distances to all the objects in the training dataset are calculated and *k* nearest of them are selected. The decision is made based on simple voting, i.e. if the majority of the selected neighbours are squares, the new object will be classified as a square and vice versa. It is reasonable to select *k* equal to some odd number to avoid the situation of equal votes.

The algorithm described above has many drawbacks. First, the choice of *k* is crucial for the classification results. In the example, if only three neighbours are considered ($k = 3$), the new object is classified as a square, but if $k = 5$, it is classified as a circle. When the number of possible classes is more than two, situations of equal votes become possible and there is a need for additional criteria to make a decision. If a certain type of object is dominating in the training set, it may also tend to "win" the votes more often just because of the denser population and higher probability of appearing near the tested object. Other algorithms, such as support vector machines (SVMs) and artificial neural networks (ANNs) often treat these problems better[28,29].

Figure 8 illustrates the classification of the speed profiles using the nearest neighbour algorithm ($k = 1$). Six profiles of each type with very typical shapes are selected as a training dataset (Fig. 8*a*) and the results of the classification are presented in Figure 8*b* (only 75 profiles are plotted to keep the diagram readable). The results appear to be more robust compared to unsupervised cluster-

a)



b)

			k-means			Total
			situation type			
			"a"	"b"	"c"	
Observer	situation type	**"a"**	***32%***	*8%*	*0%*	**40%**
		"b"	*0%*	***6%***	*0%*	**6%**
		"c"	*2%*	*9%*	***43%***	**54%**
Total			**34%**	**23%**	**43%**	**100%**

Correct classifications - *81%*

Fig. 6 Results of k-means clustering of the speed profiles (a) and comparison with the observer's results (b)

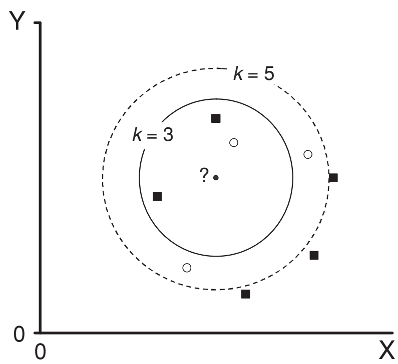


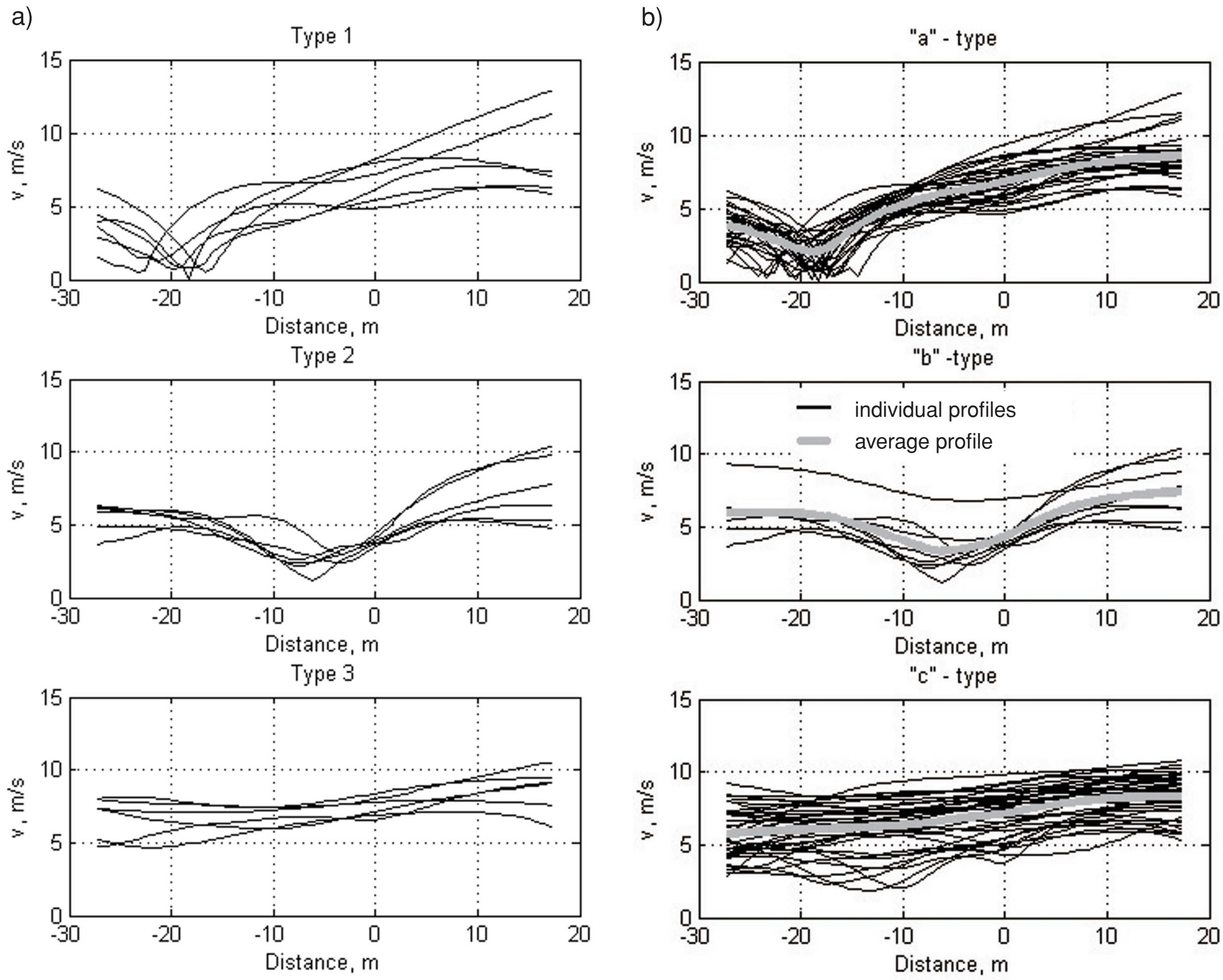
Fig. 7 *k*-nearest neighbours classification

ing, but preparation of the training set requires additional manual work.

3.4 Dimension reduction

The n-co-ordinate set gives the exact description of the vector (profile shape). However, operating all n co-ordinates is not always convenient as it complicates data visualisation and analysis, increases computation time and may even yield less accurate results compared to analysis of simplified data. A possible solution is to find an approximation to a vector that may be described by fewer co-ordinates but still preserves enough information about the vector's important features.

This problem may be solved in several possible ways. One is to select several co-ordinates from the original set, given that they contain the most important information (feature selection). Another way is to create a completely new co-ordinate system with fewer dimensions



c)

			nearest neighbour			Total
			situation type			
			"a"	*"b"*	*"c"*	
Observer	situation type	*"a"*	***36%***	*0%*	*4%*	**40%**
		"b"	*0%*	***3%***	*3%*	**6%**
		"c"	*2%*	*0%*	***52%***	**54%**
Total			**38%**	**3%**	**59%**	**100%**

Correct classifications - *91%*

Fig. 8 Training set (a), the nearest neighbour classification based on the training set (b) and comparison with the observer's results (c)

than n and project the original vector in the new space (feature extraction). The task here is to find a system that preserves as much information about the vector's features as possible and omits less important information.

To explain in a simple way how this works, let us consider a vector $\boldsymbol{a}$ in 3-dimensional space (Fig. 9). Let $\boldsymbol{e_1}, \boldsymbol{e_2}, \boldsymbol{e_3}$ be an orthonormal co-ordinate system. A complete vector description is given by its three co-ordinates, i.e. its projections on the base vectors ($\boldsymbol{a} = c_1 \cdot \boldsymbol{e_1} + c_2 \cdot \boldsymbol{e_2} + c_3 \cdot \boldsymbol{e_3}$). Omitting one of the co-ordinates, for example c_3, we approximate the original vector with its projection $\boldsymbol{a_{12}}$ on the plane defined by the base vectors $\boldsymbol{e_1}$ and $\boldsymbol{e_2}$. If we omit two co-ordinates, c_2 and c_3, the original vector is approximated with its projection $\boldsymbol{a_1}$ on the base vector $\boldsymbol{e_1}$. Obviously, as more co-ordinates are omitted, the quality of the approximation decreases. However, how much information is lost when a co-ordinate depends on the chosen co-ordinate system and which of the co-ordinates are omitted. For example, if the orientation of the vector is close to one of the base vectors, this co-ordinate will be most important. Omitting it will imply that nearly all the information about the vector is lost. Other co-ordinates, on the contrary, may be easily omitted without introducing any substantial errors in vector description.

One of the feature extraction techniques is singular value decomposition (SVD[30]). Let us construct a matrix M in which each m column represents a vector of length n. According to the SVD theorem a $n \times m$ matrix M can be presented in the form of a product of three components:

$$M = U \cdot S \cdot V^T,$$

where U and V are unitary matrices of $n \times n$ and $m \times m$ size respectively and S is a $n \times m$ diagonal matrix with non-negative values on the diagonal arranged in a non-increasing order.

One of the interpretations of the SVD results is that U is a set of orthonormal vectors defining a new co-ordinate system, while $S \cdot V^T$ are the co-ordinates of the original vectors in the new co-ordinate system. An important property of SVD is that the new base vectors are sorted in decreasing order of their importance, i.e. the first co-ordinate gives more information about the original vectors than the second one and so on. Another important property is that if one wants to use only i co-ordinates to describe the original vector set ($i < n$), the first i base vectors are always the best possible (in a mean square sense) co-ordinate system to describe the original vectors. The calculation of SVD is quite a common method in linear algebra and is often implemented in specialised software (e.g. Matlab) as a standard function.

The singular value decomposition was applied to the matrix 60×253 constructed from the speed profile vectors' co-ordinates. The new co-ordinate system U contained 60 base vectors and each profile was described by 60 co-ordinates contained by the $S \cdot V^T$ matrix. However, the co-ordinates decreased rapidly from the first towards the last one, which proved that the most characteristic information on the speed profile was contained by the first few co-ordinates. Only two first co-ordinates were chosen as an approximation.

The first 25 speed profiles were manually sorted into three categories. Figure 10 shows a two-dimensional plot where each profile is presented by a point defined by the two co-ordinates. It is clearly seen that the three profile types form quite distinctive clusters in the plot space.

After the number of dimensions is decreased, the clustering or supervised learning techniques may be used to distinguish the profile shapes. The profiles in Figure 11 are split using a simple threshold criteria set for the two co-ordinate values (again, only 75 profiles are plotted).

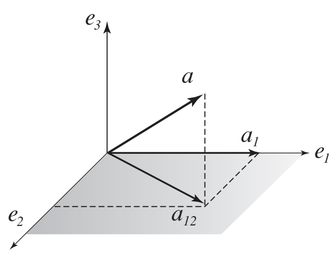


Fig. 9 Approximations of a vector in 3-dimensional space

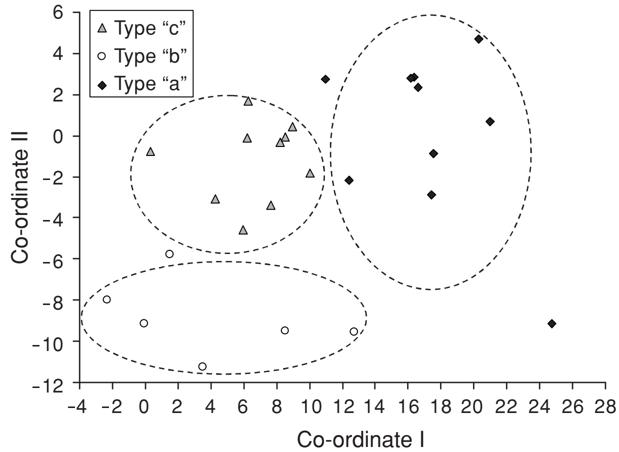
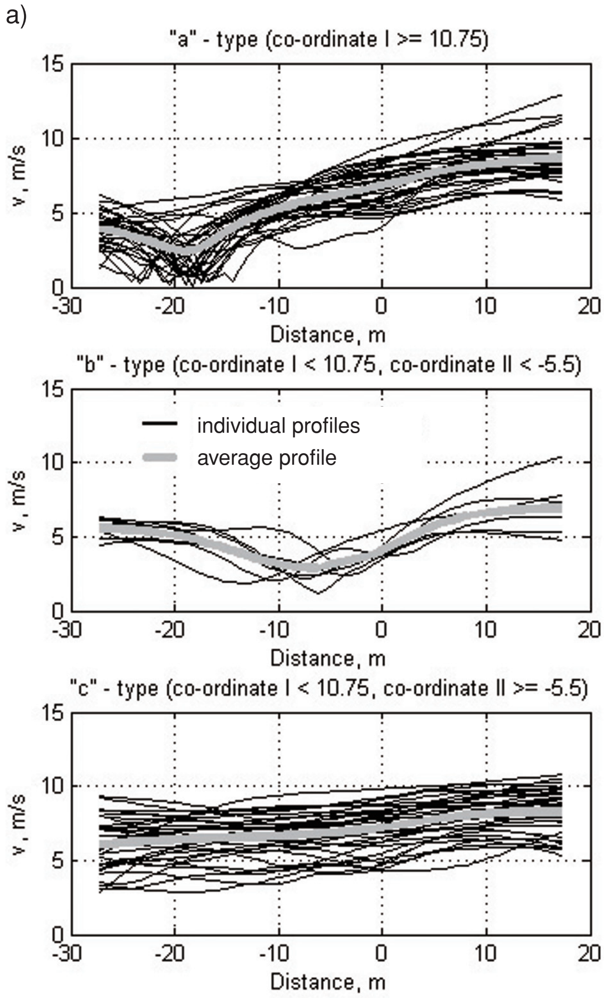


Fig. 10 The manual classification of the speed profiles presented by their 2-dimensional approximations

a)



b)

			SVD situation type			Total
			"a"	*"b"*	*"c"*	
Observer	situation type	*"a"*	***36%***	*0%*	*4%*	**40%**
		"b"	*0%*	***3%***	*3%*	**6%**
		"c"	*8%*	*0%*	***46%***	**54%**
Total			**44%**	**3%**	**53%**	**100%**

Correct classifications - *85%*

Fig. 11 The results of profile classification based on 2-dimensional approximations (a) and comparison with the observer's results (c)

The classification accuracy of the described technique is dependent on the frequency of the profiles of different shapes. Thus, if a certain type is clearly dominant, it will also "draw" the new co-ordinate system to its side, which may result in the fact that the features of the other, less frequent types, will be lost to a greater extent. Other techniques for dimension reduction and data visualisation that may be tested are, for example, multidimensional scaling (MS)[31] and isomap[32].

4. DISCUSSION AND CONCLUSIONS

There is a great diversity in the road user behaviour forms and classifying it is not a simple task, neither for a human, nor for an automated technique. In this paper, we evaluate three pattern recognition techniques according to three criteria: I – quality of the differentiation between the behaviour types; II – use of the variation in the data attributable to different behaviour (profile shapes) and III – ability to treat large datasets.

Criteria II and III are fulfilled by all three methods. The methods utilise the information about profile shapes and can be applied to extensive datasets. When it comes to criterion I, the quality of the results produced by the three illustrated techniques are not the same.

The k-means clustering, which is the least successful, but the simplest one to implement and test, can be recommended when the results have to be produced quickly. The clustering algorithms often assume that there is a set number of clusters in the data and try to "force" the data into them. If the number of clusters is high, but the data are quite homogeneous, the clusters will most probably look alike and the differences might be difficult to interpret. On the other hand, using a higher number of clusters may reveal small peculiar groups not obviously seen among the other profile shapes that are more common. Moreover, more advanced algorithms are capable of making a decision on the optimal number of clusters for the current dataset.

The supervised learning and classification based on the nearest neighbour algorithm shows the best results. The training set reflects human judgements of the expected number of groups and the typical shapes the classification is based on, i.e. it is exogenous to the data. If a group is missed at the initial stage of classification, it will not be represented in the training set and will be missed in the final classification as well. Another problem is that production of an extensive training set requires much of the manual input.

The dimension reduction techniques are more complex to implement and based on extracting the most typical features from the data and operating the features rather than the original data. This simplifies the analysis and data visualisation to a great extent. For example, drawing up the 2-dimensional approximations of the profiles produced by the SVD algorithm (Fig. 10) reveals the pattern of location of different profile types on the graph and allows the setting up of very simple classification criteria that still perform very well. There is a risk, however, that the omitted dimensions may contain important information for the final classification of the profiles and thus affect the quality of the results. As for the quality of the classification, it is still lower compared to supervised learning.

A problem that needs special investigation is the profiles with shapes that do not match any of the typical patterns. All three techniques are quite insensitive to such outliers and simply force them into one of the typical groups. However, examination of the outliers might be important in case they represent some kind of breakdown in normal traffic that might have implication for safety or efficiency. Detailed examination of such situations might give an idea on how they can be eliminated. A possible solution can be to compare individual profiles with the average profile and select significantly different ones.

Finding the right technique for the data is often stated to be more an art than a science, and parameters working well for one dataset will not certainly work for another. The best strategy in this case is to have a toolbox of different techniques where the right one is found by using trials.

The general conclusion is that the pattern recognition techniques perform quite well in classifying the behaviour types compared to classification by a human observer. The advantage of these techniques is the automation of the classification process which allows for analysing large datasets. Another aspect is the reduction of the subjective effects a specific observer might have on the results when doing the classification manually. We have argued in principle that the subjective component introduced by an observer might be useful, especially if the differences in behaviour are difficult to express in objective terms. On the other hand, pattern recognition techniques might help reveal the relations between this subjective dimension and objective variables, contribute to standardisation and therefore allow for larger comparability between analyses made by different individuals.

As a direction for future work, pattern recognition techniques can be tested in classifying more complex situations that a human observer manages to classify without being able to explicitly formulate the classification criteria. For example, there is a great interest in finding objective parameters that better reflect the severity of traffic conflicts. This, however, requires a large set of traffic conflicts with detailed description of the road users movements to be available.

REFERENCES

1. Englund, A., N.P. Gregersen, C. Hydén, P. Lövsund and L. Åberg. Trafiksäkerhet – en kunskapsöversikt (*in Swedish*) Traffic safety - a knowledge review. KFB, Studentlitteratur, Lund. (1998).
2. Aarts, L. and I. van Schagen. Driving speed and the risk of road crashes: A review. "Accident Analysis and Prevention" 38: pp. 215-224. (2006).
3. Kloeden, C.N., A.J. McLean, V.M. Moore and G. Ponte. Travelling Speed and the Risk of Crash Involvement. Volume I: findings. University of Adelaide, Road Accident Research Unit. Report No. CR 172. (1997).
4. Nilsson, G. Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety. Doctoral thesis. Lund University, Department of Technology and Society, Traffic Engineering. (2004).
5. Ericsson, E. Urban driving patterns - characterisation, variability and environmental implications. Doctoral thesis. Lund University, Lund Institute of Technology, Department of Technology and Society. (2000).
6. Hagring, O. Framkomlighet i korsningar utan trafiksignaler, en litteraturöversikt (*in Swedish*) Traffic efficiency of non-signalised intersections, a literature review. Lunds Tekniska Högskola, Institutionen för Teknik och Samhälle, Lunds Universitet. Bulletin 190. (2000).
7. Nygård, M. A method for analysing traffic safety with help of speed profiles. Master thesis. Tampere University of Technology, Department of Civil Engineering. (1999).
8. Malkhamah, S., M. Tight and F. Montgomery. The development of an automatic method of safety monitoring at Pelican crossings. "Accident Analysis and Prevention" 37: pp. 938-946. (2005).
9. Larsson, H. Förarstöd för lägre bränsleförbrukning och minskade emissioner: utvärdering av två system (*in Swedish*) Driver support for better fuel consumption and lower emissions: evaluation of two systems. Licenciate thesis. Lund University, Department for Technology and Society, Faculty of Engineering LTH. (2009).
10. Hydén, C. and A. Várhelyi. The effects on safety, time consumption and environment of large scale use of roundabouts in an urban area: a case study. "Accident Analysis and Prevention" 32: pp. 11-23. (2000).
11. Pau, M. and S. Angius. Do speed bumps really decrease traffic speed? An Italian experience. "Accident Analysis & Prevention" 33: pp. 585-597. (2001).
12. Karlgren, J. Bilars hastighet längs gator med gupp (*in Swedish*) Vehicles speed along streets with humps. Doktorsavhandling. Chalmers Tekniska Högskola, Göteborg, Tema Stad & Trafik, Sektionen för Arkitektur. (2001).
13. Michon, J.A. A critical view of driver behaviour models: what do we know what should we know? in L. Evand and R. C. Schwing (Eds.). Human behaviour and traffic safety. Plenum Press, New York. (1985).
14. Peräaho, M., E. Keskinen and M. Hatakka. Driver competence in a hierarchial perspective: implications for driver education. University of Turku, Traffic Research. Report for Swedish Road Administration. (2003).
15. Carsten, O.M.J., D.J. Sherborne and J.A. Rothengatter. Intelligent traffic signals for pedestrians: evaluation of trials in three countries. "Transportation Research Part C" 6: pp. 213-229. (1998).

16. Hakkert, A., V. Gitelman and E. Ben-Shabat. An evaluation of crosswalk warning systems: effects on pedestrian and vehicle behaviour. "Transportation Research Part F" 5: pp. 275-292. (2002).
17. Martinez, K.L.H. and B.E. Porter. Characterizing red light runners following implementation of a photo enforcement program. "Accident Analysis and Prevention" 38: pp. 862–870. (2006).
18. Grayson, G.B. The Malmö study: a calibration of traffic conflict techniques. SWOV. Report R-84-12. (1984).
19. Hydén, C. The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis. Lund University, Department of Traffic Planning and Engineering. (1987).
20. Ericsson, E. Variability in urban driving patterns. "Transportation Research Part D" 5: pp. 337-354. (2000).
21. André, M., J. Hickman, D. Hassel and R. Jourmard. Driving cycles for emission measurements under European conditions. SAE. Technical Paper 950926. (1995).
22. Laureshyn, A., H. Ardö, T. Jonsson and Å. Svensson. Application of automated video analysis for behavioural studies: concept and experience. "IET Intelligent Transport Systems" 3(3): pp. 345-357. (2009).
23. Parkhurst, D. Using digital video analysis to monitor driver behaviour at intersections. Center for Transportation Research and Education, Iowa State University. CTRE Project 05-214. (2006).
24. Messelodi, S., C.M. Modena and M. Zanin. A computer vision system for the detection of vehicles at urbam road intersections. Instituto Trentino di Cultura (ITC). Technical Report T04-02-07. (2004).
25. Várhelyi, A. Driver's speed behaviour at a zebra crossing: a case study. "Accident Analysis and Prevention" 30(6): pp. 731-743. (1998).
26. Sekine, T. and E. Sekine. Characteristic classification of driving behaviour with mixed traffic situation for driver assist. 21st International Symposium on Dynamics of Vehicles on Roads and Tracks, Stockholm, 17-21 August. (2009).
27. Ardö, H. Multi-target tracking using on-line viterbi optimisation and stochastic modelling. Doctoral thesis. Lund University, Faculty of Engineering LTH, Centre for Mathematical Science. (2009).
28. Duda, R.O. and P.E. Hart. Pattern classification and scene analysis. A Wiley-interscience publication, New York. (1973).
29. Ripley, B.D. Pattern recognition and neural networks. Cambridge University Press. (1996).
30. Strang, G. Introduction to applied mathematics. Wellesley-Cambridge Press, Wellesley, Massachusetts. (1986).
31. Borg, I. and P. Groenen. Modern Multidimensional Scaling: theory and applications. Springer-Verlag, New York. (2005).
32. Tenenbaum, J.B., V. de Silva and J.C. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction. "Science" 290 (5500): pp. 2319-2323. (2000).